### CPEN 311: Digital Systems Design

Circuit Timing: Part 2: Practical Issues

### Learning Objectives

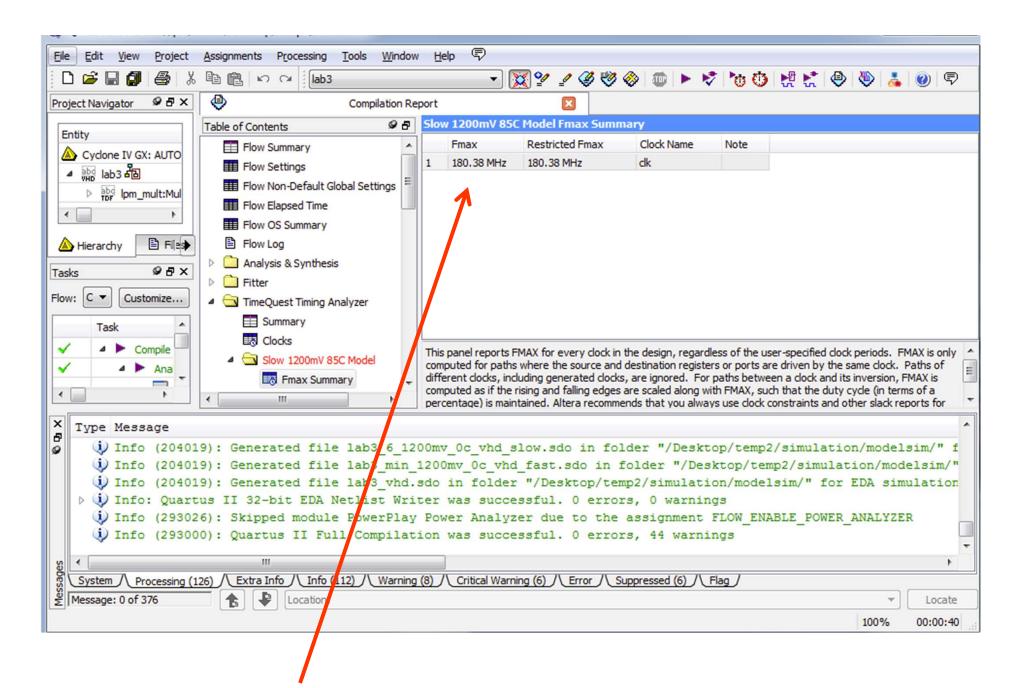
- 1. Understand what timing closure is and why it is difficult
- 2. Understand how pipelining can help with timing closure
- 3. Understand retiming and be able to apply it to a circuit
- 4. Be able to discuss the effects of clock skew
- 5. Understand what a PLL is used for
- 6. Understand the cause and impact of glitches caused by unequal combinational path delays

### Timing Closure

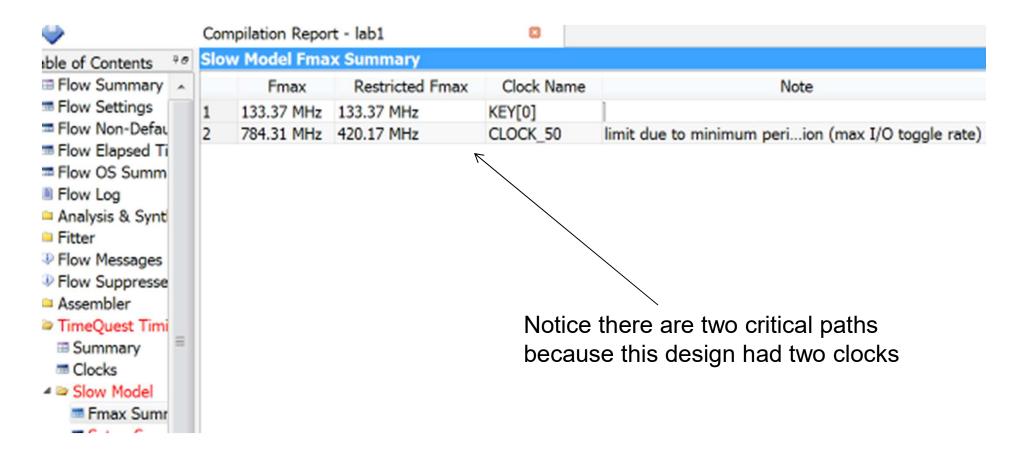
**Timing Closure**: Ensuring your design meets timing constraints

So far you have been compiling and just hoping it runs at the required speed (eg. If you are using CLOCK\_50, the critical path <= 20 ns)

That was fine for these simple labs, but for complex designs, a simple compile may not lead to you meet your timing constraints.



#### From my implementation of the lab:

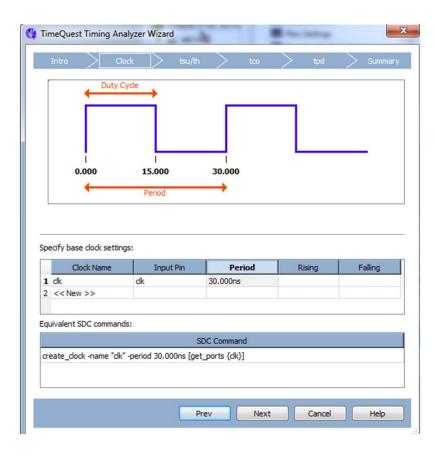


Fmax means the maximum frequency the circuit can run at

### Timing Constraints in Quartus II

# You can specify a desired target clock frequency aka Timing Constraint

Assignments Menu → Timing Analysis Wizard



### **Quartus Timing Optimization**

- Tries to meet your desired clock frequency
  - By default, an unachievable value (eg, 1GHz)
  - Better to specify actual target (eg, 50MHz)
- May not meet target
  - Limited by logic + routing delays in FPGA

When you ask for...

#### slow clock frequency

smaller circuits → optimize total number of logic gates

#### fast clock frequency

larger circuit → optimize number of logic gates along path

All about tradeoffs!

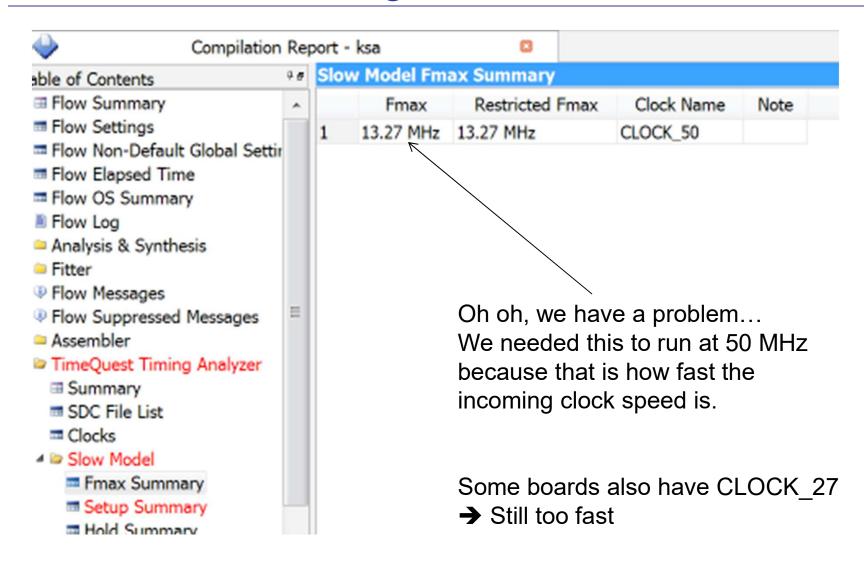
### Tradeoffs: Lab 1

#### Experiments conducted on a DE2 board:

Scenario	Achieved Fmax	Area of Circuit
No explicit timing constraints (1 GHz)	133 MHz	148 Logic Elements
Timing constraint of 130 MHz	130 MHz	143 Logic Elements
Timing constraint of 10 MHz	112 MHz	135 Logic Elements

The next few slides are meant as a kind of "case study" so you understand what timing closure is all about.

### **Timing Violations**



### Let's investigate: Worst-Case Timing Paths

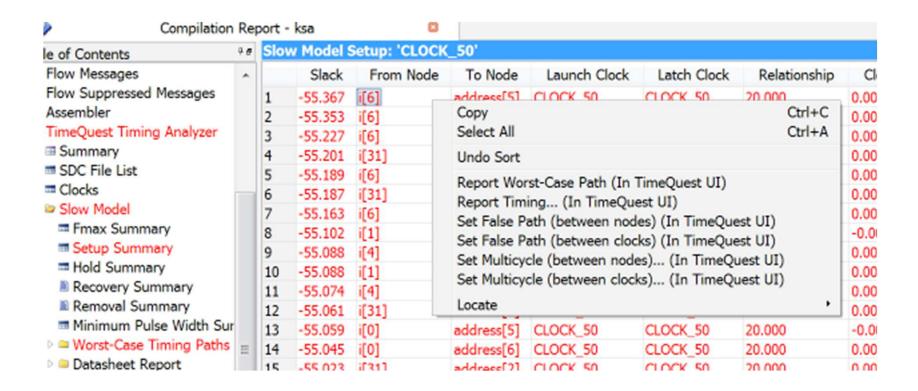
Compilation	· ksa	0						
Table of Contents	40	Slo	w Model S	Setup: 'CLOCK	_50'			
Analysis & Synthesis	*		Slack	From Node	To Node	Launch Clock	Latch Clock	Relationship
□ Fitter		1	-55.367	i[6]	address[5]	CLOCK 50	CLOCK_50	20.000
Flow Messages		2	-55.353	i[6]	address[6]	CLOCK 50	CLOCK 50	20.000
Flow Suppressed Messages		3	-55.227	i[6]	address[3]	CLOCK 50	CLOCK 50	20.000
Assembler		4	-55.201	i[31]	address[5]	CLOCK 50	CLOCK 50	20.000
TimeQuest Timing Analyzer	≡ Ir	5	-55.189	i[6]	address[2]	CLOCK 50	CLOCK_50	20.000
Summary		6	-55.187	i[31]	address[6]	CLOCK 50	CLOCK 50	20.000
SDC File List		7	-55.163	i[6]	address[1]	CLOCK 50	CLOCK 50	20.000
□ Clocks		8	-55.102	i[1]	address[5]	CLOCK 50	CLOCK 50	20.000
■ Slow Model		9	-55.088	i[4]	address[5]	CLOCK 50	CLOCK_50	20.000
Fmax Summary		10	-55.088	i[1]	address[6]	CLOCK 50	CLOCK 50	20.000
■ Setup Summary		11	-55.074	i[4]	address[6]	CLOCK 50	CLOCK 50	20.000
■ Hold Summary		12	-55.061	i[31]	address[3]	CLOCK 50	CLOCK 50	20.000
Recovery Summary		13	-55.059	i[0]	address[5]	CLOCK_50	CLOCK_50	20.000
Removal Summary		14	-55.045	i[0]	address[6]	CLOCK 50	CLOCK 50	20.000
Minimum Pulse Width Sur		15	-55.023	i[31]	address[2]	CLOCK 50	CLOCK 50	20.000
Worst-Case Timing Paths		16	-55.010	i[6]	address[4]	CLOCK 50	CLOCK 50	20.000
Setup: 'CLOCK_50'		17	-54.997	i[31]	address[1]	CLOCK_50	CLOCK_50	20.000

Slowest path is from bit 6 of the i register to bit 5 of the address input of s\_mem

"Slack" is difference between "this path delay" and "target clock period" (20ns). Negative slack is BAD. Often hear about "worst-case negative slack" or WCNS<sub>11</sub>

### What is this path?

#### Right click, Report-Worse-Case Path:



Can go into Timing Quest and select Create Timing Network and then come back...

ath Summary	Statis	stics	Data P	ath Wa	eveform	
ta Arrival Pa	th					
Total	Incr	RF	Type	Fanout	Location	Element
0.000	0.000					launch edge time
4 2.755	2.755					clock path
2.755	2.755	R				clock network delay
4 78.165	75.410					data path
3.005	0.250		uTco	1	LCFF_X42_Y25_N17	i[6]
3.005	0.000	RR	CELL	7	LCFF_X42_Y25_N17	i[6] regout
3.527	0.522	RR	IC	1	LCCOMB_X41_Y25_N26	Mod0 auto_generated divbs_num cs1a[6]~21 dataa
3.940	0.413	RR	CELL	1	LCCOMB_X41_Y25_N26	Mod0 auto_generated dinum cs1a[6]~21 combout
4.215	0.275	RR	IC	1	LCCOMB_X41_Y25_N2	Mod0 auto_generated divbs_num cs1a[6]~22 datab
4.608	0.393	RR	CELL	4	LCCOMB_X41_Y25_N2	Mod0 auto_generated dinum cs1a[6]~22 combout
4.870	0.262	RR	IC	1	LCCOMB_X41_Y25_N16	Mod0 auto_generated divbs_num cs1a[8]~23 datad

It is coming out of register i and going to a divider...

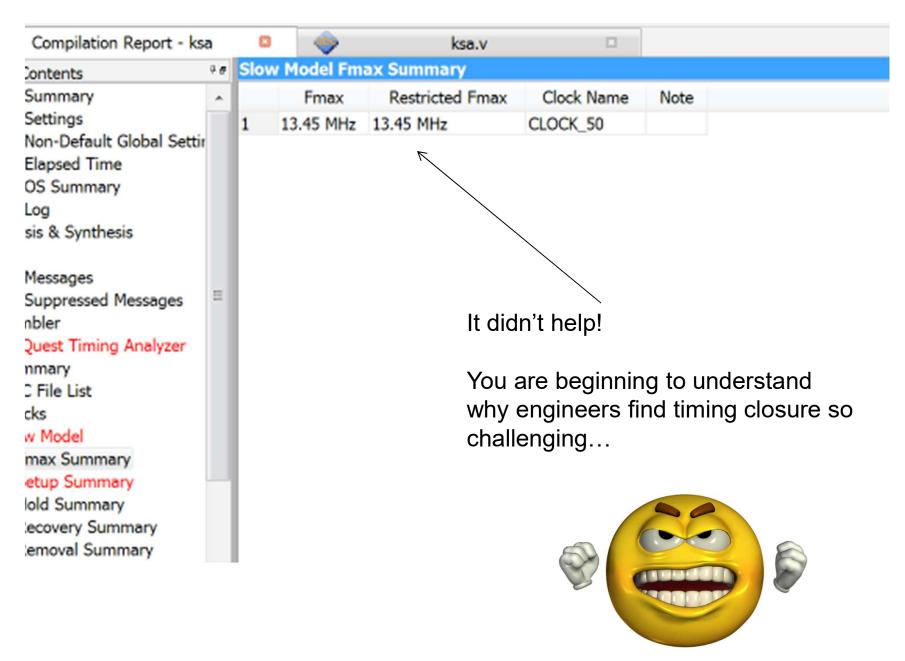
Combinational dividers are notoriously slow...

Look at the code to see what is happening...

In my code, I found something like:

I suspect the division might be due to this mod function. But, if course, mod 256 is really just taking the 8 order LSB. We could get rid of that and just remove the % 256.

So, I do it and recompile...



Going back to my code, I see:

Certainly if we are dividing by 3, this should be a small enough circuit right?

But wait... in my implementation, i is an integer (32 bits) and the constant 3 is interpreted as 32 bits. So this constructs a 32 bit divider!

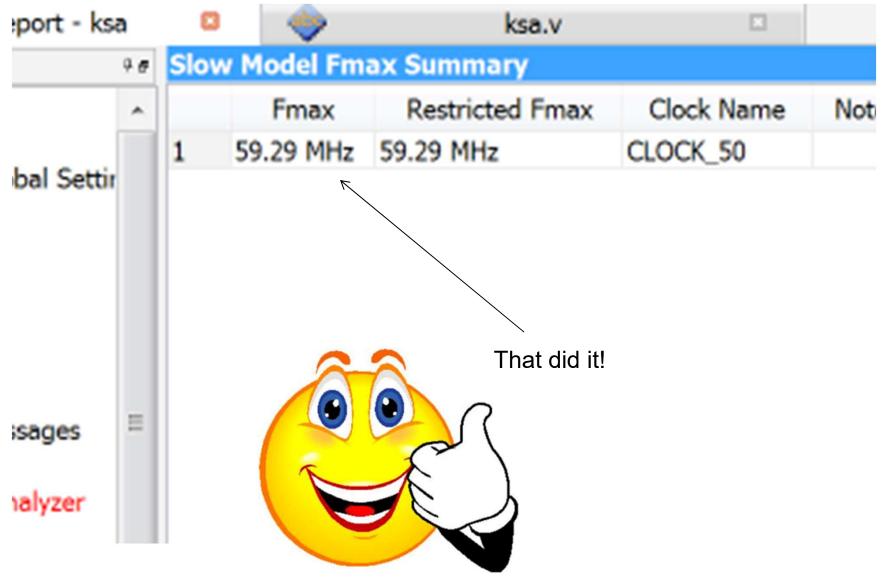


#### So, replace it with this:

```
state_swap_3: begin
    oldsi = q; // s[i] is now available
    j = ( j + q + secret_key[i%3]);

state_swap_3: begin
    oldsi = q; // s[i] is now available
    j = ( j + q + secret_key[i[7:0]%2'b11]);
    // read s[j]
```

Then recompile and ...



That was a case study to introduce you to the frustrations concept of timing closure.

#### Can Go Faster Than Predicted Fmax?

#### **Overclocking**

- Quartus gives conservative estimate
- Actual delays vary chip-to-chip

#### Dangerous... don't do it

In testing, it might <u>appear</u> to work.
 But, how do you know that you are exercising the critical path?

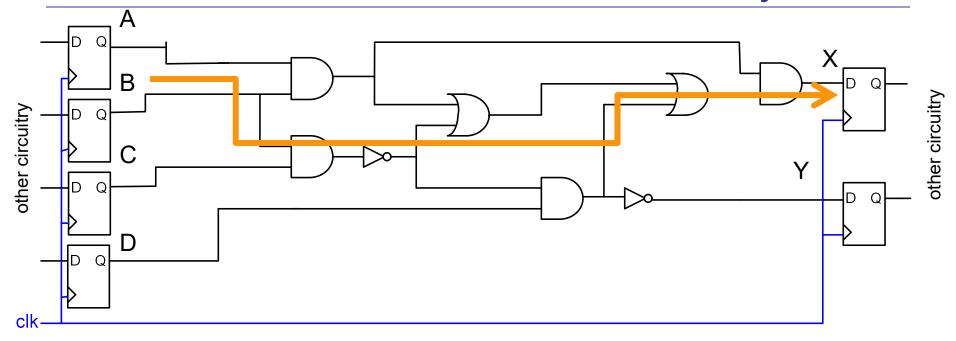
Some chips will be slower than the one used for prototyping

 Conditions such as temperature can affect gate delays



# **PIPELINING**

### From earlier: Critical Path Delay

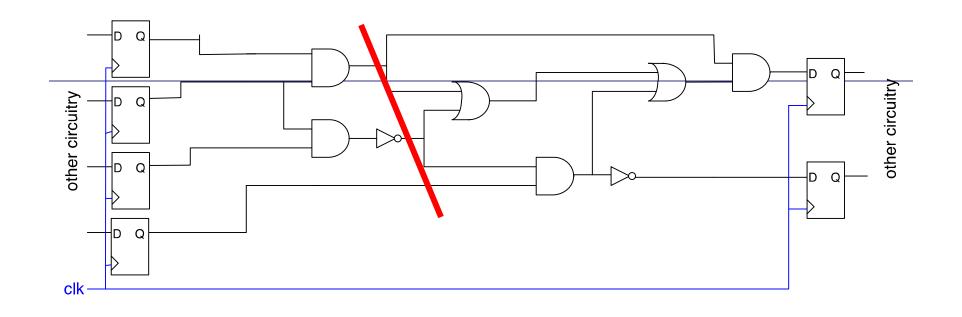


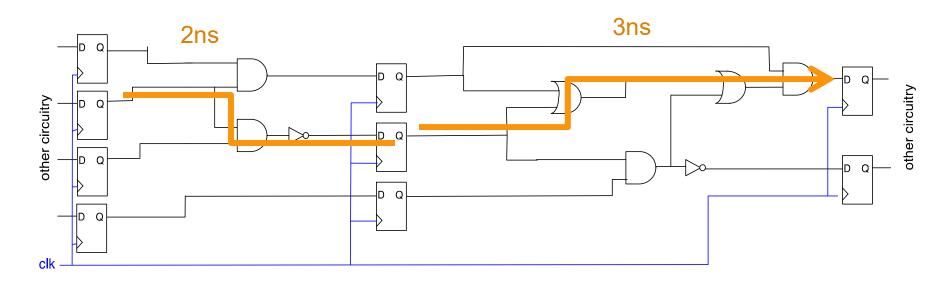
**Critical Path is** 

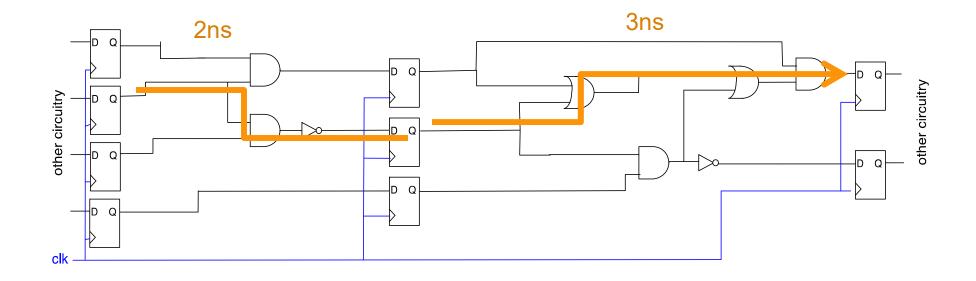
 $B \rightarrow AND \rightarrow INV \rightarrow OR \rightarrow OR \rightarrow AND \rightarrow X$ 

Critical Path Delay is 5ns

Clock period cannot be smaller than 5ns or else X register will read in wrong data







The new longest path between registers is 3 ns.

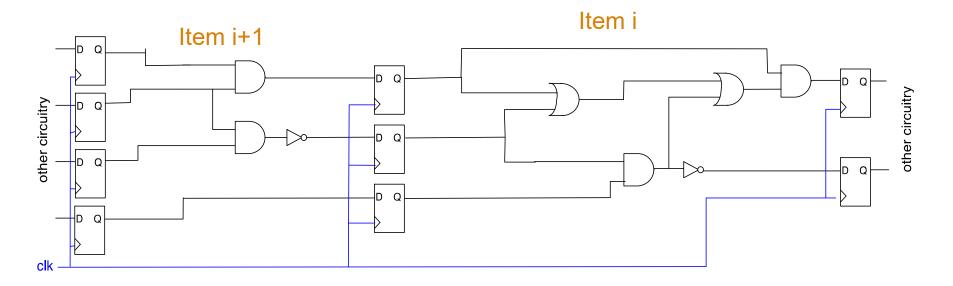
So, this now goes at 1/3ns = 333 MHz (before it was 1/5 = 200 MHz)

But it takes 2 cycles to get a result now.

Have we really improved anything?

#### Pipelining might help if:

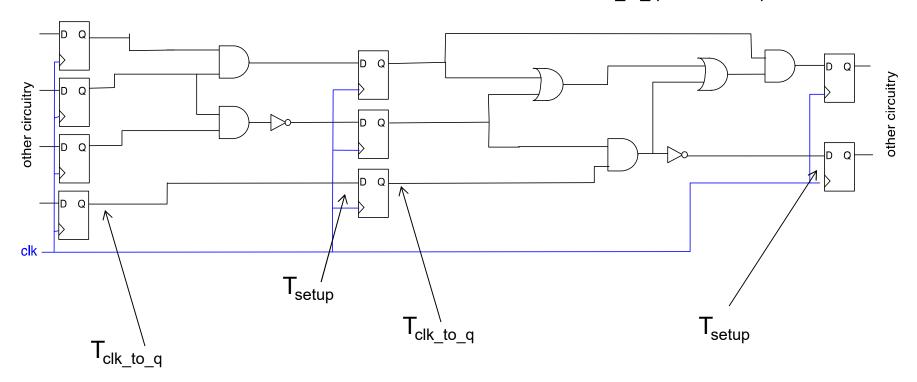
- One (or several) of the paths are very long, compared to the others.
   In that case, these slow paths dictate performance. If we pipeline the slow parts, we can run the whole circuit faster.
- 2. We can sometimes have multiple data items "in flight" at once:



### Limits to pipelining

There is a limit to how much benefit you can get from pipelining:

- Every pipeline stage has the overhead of  $t_{\text{clk\_to\_q}}$  and  $t_{\text{setup}}$ .



In an FPGA it can be even worse.

In some FPGAs, signal must pass through extra interconnect and a LUT just to use the flip-flop.

#### Pipelining changes the timing behaviour of your circuit

 Need to take this into account when you are designing your datapath + controller

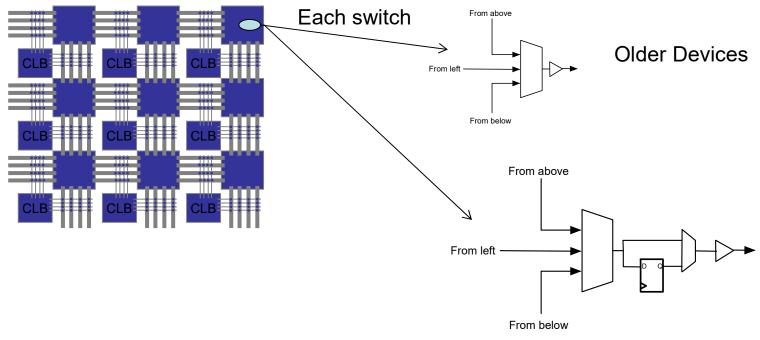
Quartus II will **not** pipeline your circuit for you. Why?

- Synthesis tools are "cycle accurate": the behaviour of the circuit must be the same, cycle-by-cycle, as the Verilog specification.
- Pipelining a design would create a different cycle-by-cycle behaviour

### Hyper-Pipelining in Stratix 10

Most recent device family has an optional flip-flop at <u>every switch</u> in the routing.

NOTE: can't use reset on these FFs (design should avoid use of resets)



Devices supporting Hyper-Pipelining (Stratix 10)

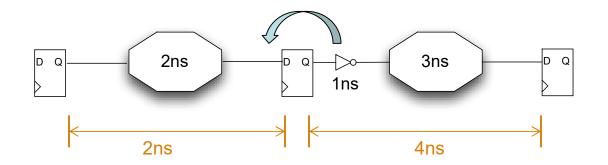
**Important Point**: Modern FPGAs support very fast clock speeds and heavily pipelined architectures. Lots of registers.

# **RETIMING**

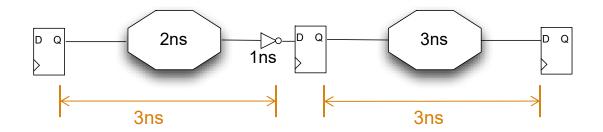
# **Retiming**

Key to pipelining: make every stage balanced (same delay)

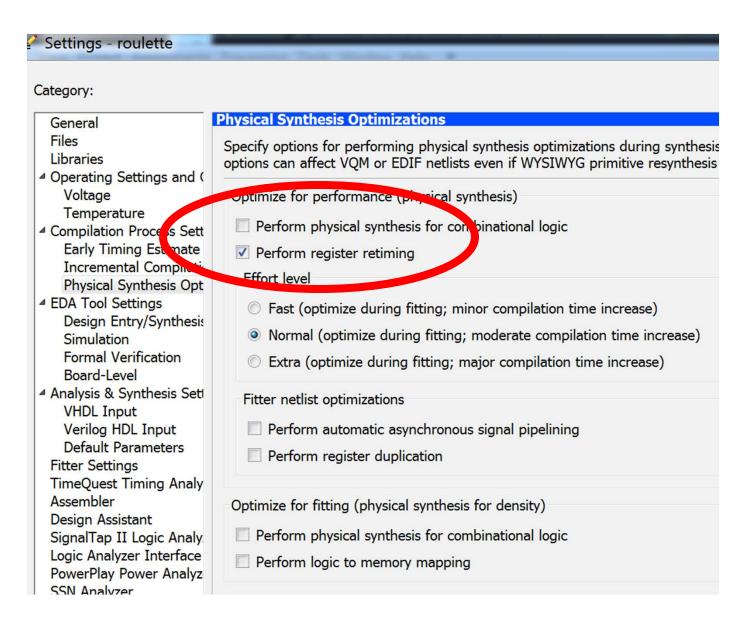
To balance, may need to move gates:



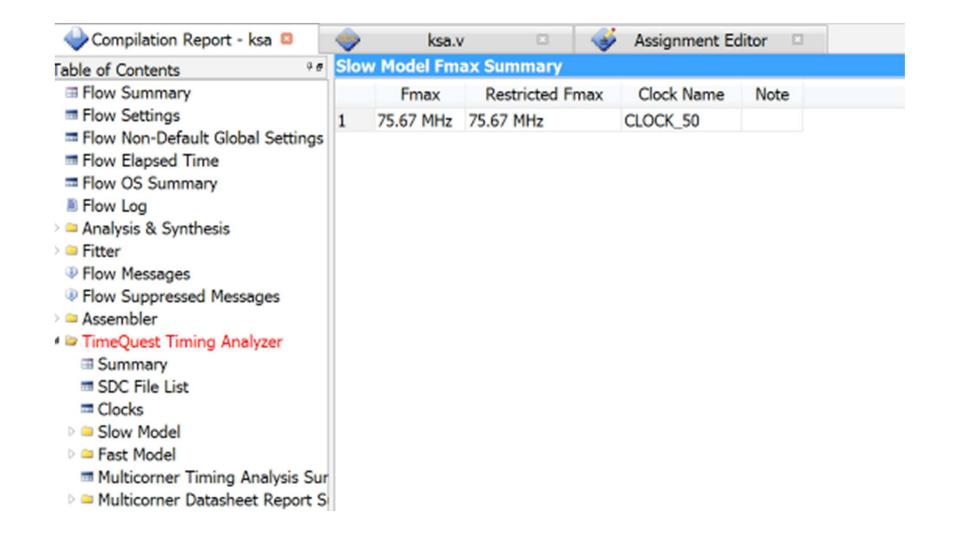
Max Freq = 1/4ns = 250 MHz



Max Freq = 1/3ns = 333 Mhz

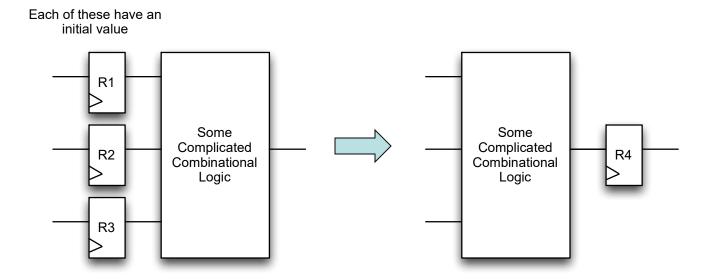


This increased my Fmax from 59 MHz to 76 MHz !! (...not always successful...)



### Initial Value Problem with Re-timing:

#### Forward Re-Timing:



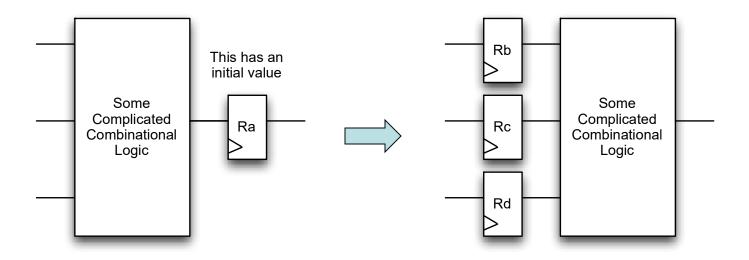
What is the initial value of R4?

Can Quartus determine it automatically?

Yes, eg simulate combinational logic with initial values of R1,R2,R3.

### Initial Value Problem with Re-timing:

#### Backward Re-Timing:

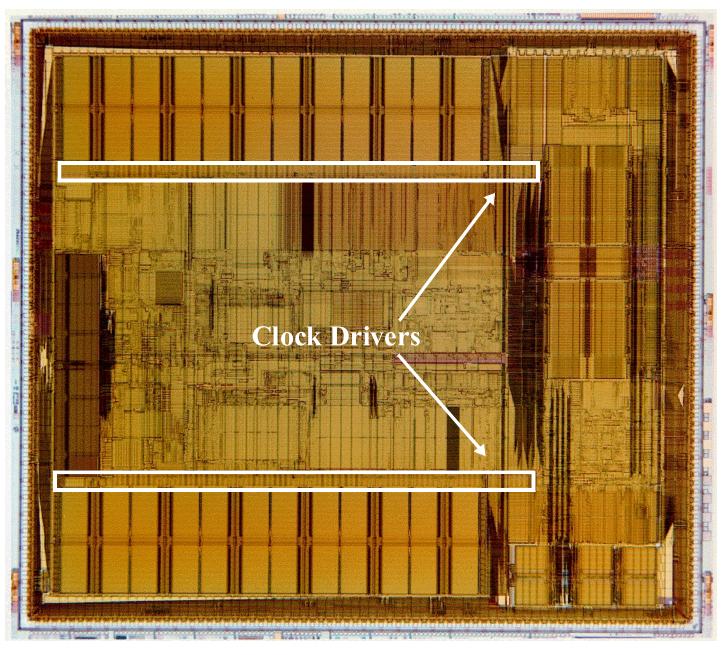


What is initial value of Rb, Rc, and Rd? (given the initial value of Ra specified by the user)

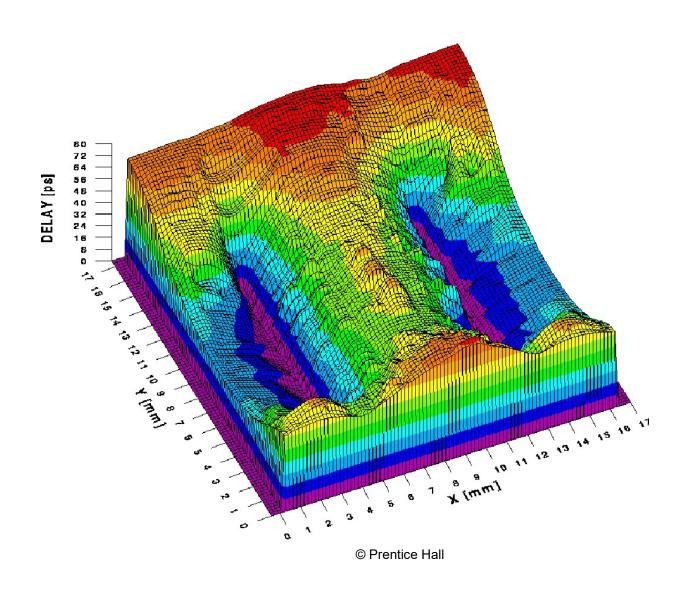
No. Simulation doesn't work "backwards". Must go through all combinations of Rb, Rc, and Rd to determine which combination gives Ra.

So: Initial values for registers may limit opportunities for retiming.

# **Clock Skew**



# Clock Skew in Alpha Processor



## Clock Skew

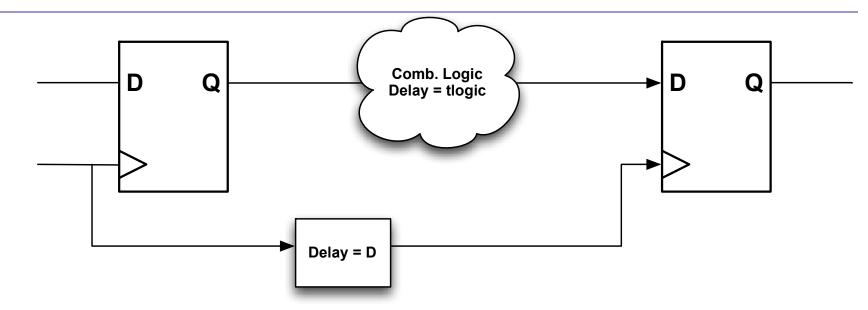
#### Clock Skew is very real:

- We can not guarantee that the clock edge arrives at all flip-flops at the same time

#### Implications:

- Improvement in Fmax (very unlikely)
- Reduction in Fmax
- Failure of the design, regardless of Fmax

## Clock Skew



If D is 0 (no skew), we know from the previous slide set that:

$$t_{clock} >= t_{clk\_to\_q} + t_{logic} + t_{setup}$$

What if D > 0 (clock skew)?

$$t_{clock} + D >= t_{clk\_to\_q} + t_{logic} + t_{setup}$$

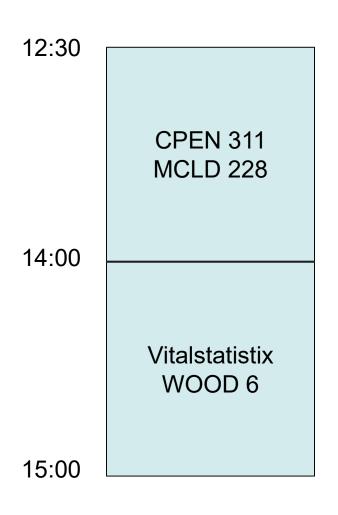
Or:

$$t_{clock} >= t_{clk\_to\_q} + t_{logic} + t_{setup} - D$$

In this case, clock skew increased our clock frequency!

# Analogy

Suppose this is your course schedule:



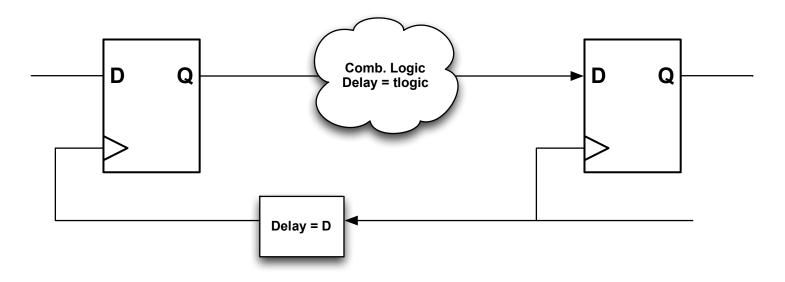
Nominally, you have **10 minutes** to get from one class to the other.

But, if next prof's watch is 2 minutes late, you have **12 minutes** to get there.

What if next prof's watch is 2 minutes early?



## Clock Skew

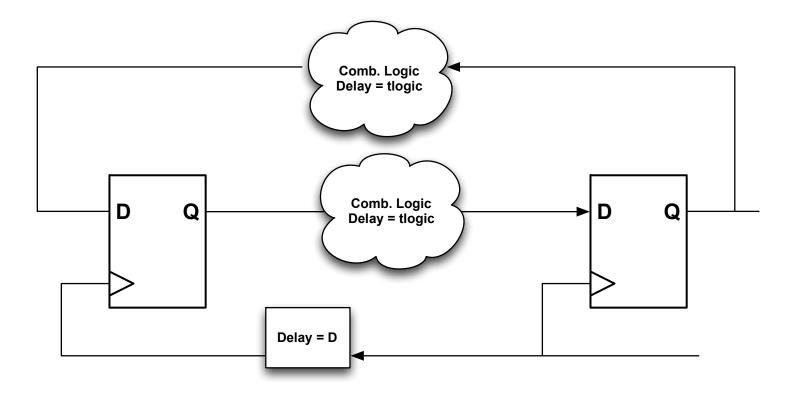


$$t_{clock}$$
 - D >=  $t_{clk\_to\_q}$  +  $t_{logic}$  +  $t_{setup}$ 

Or:

$$t_{clock} >= t_{clk\_to\_q} + t_{logic} + t_{setup} + D$$

In this case, clock skew decreased our clock frequency!



In this case, the minimum clock period:

- decreases due to top path
- · increases due to bottom path

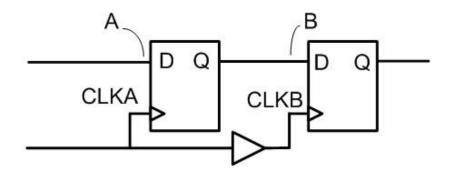
Since the critical path is the worst case path:

$$t_{clock} >= t_{clk\_to\_q} + t_{logic} + t_{setup} + D$$

ie. Overall, the clock skew increases the critical path (slows us down)

### Clock Skew

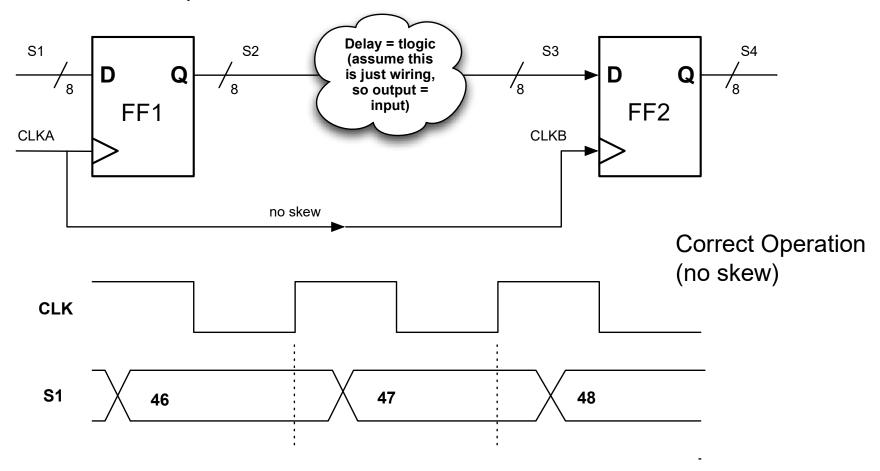
Clock skew can also cause hold time violations:



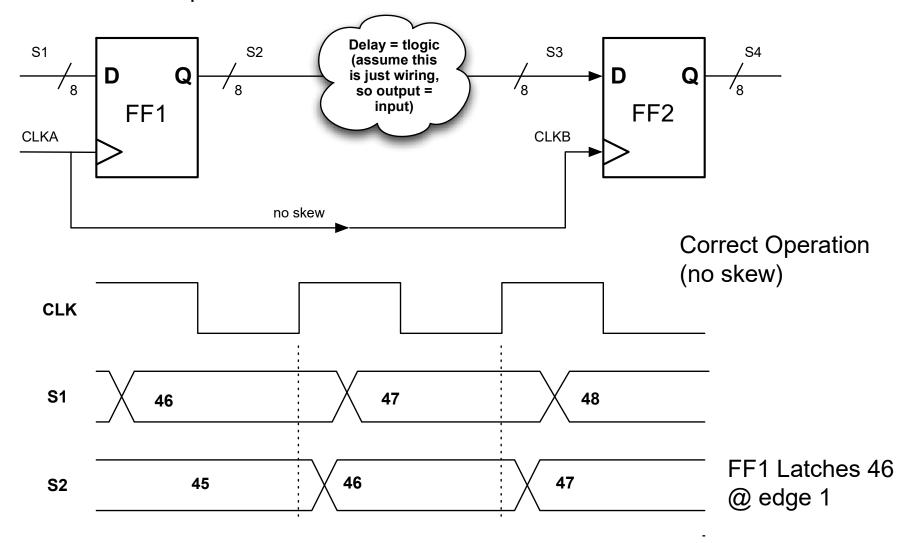
Tool needs to check for the worst case skew when checking hold time violations.

To fix, it increases routing delay to B.

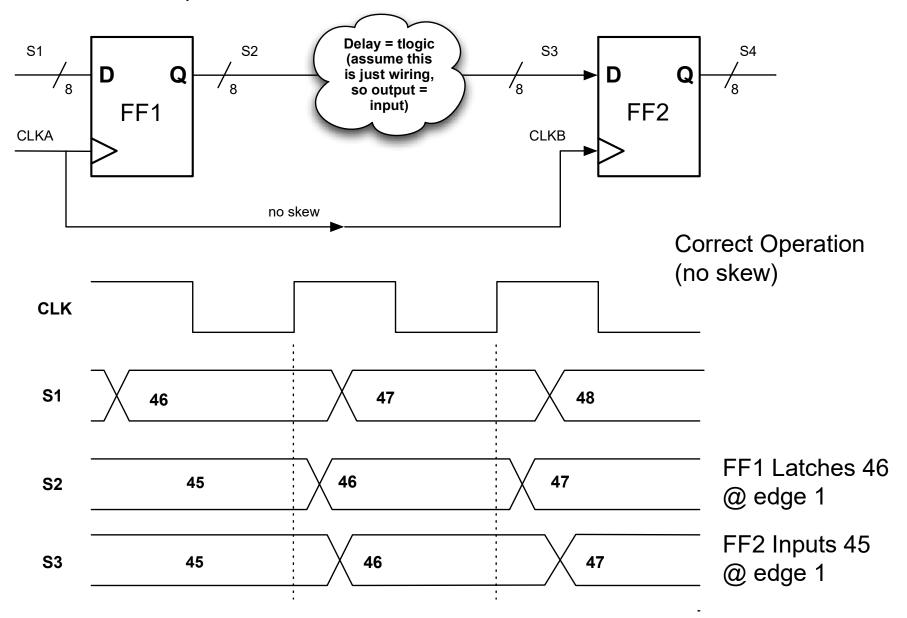
suppose B arrives *just* late enough to avoid a hold time violation what if CLKB is *delayed* with respect to CLKA?



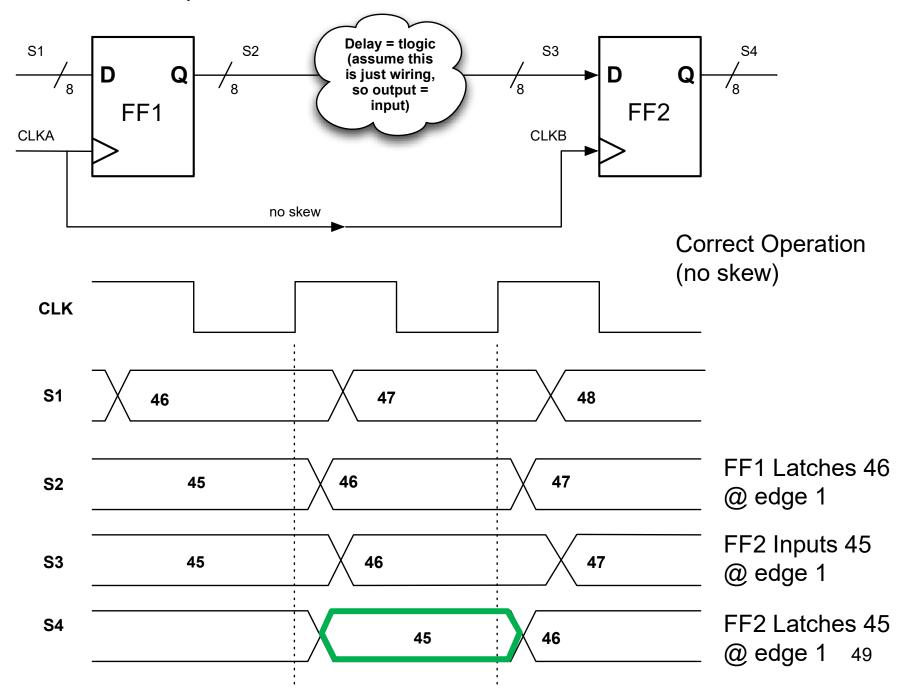
46

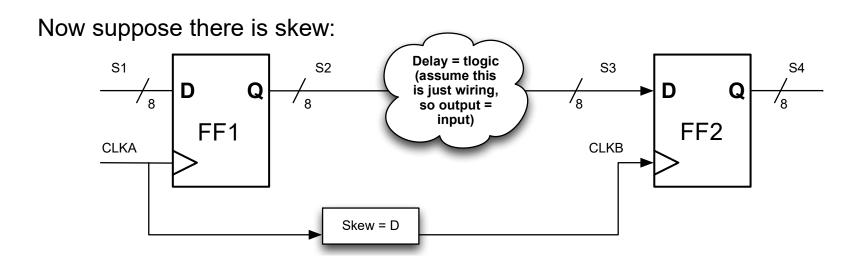


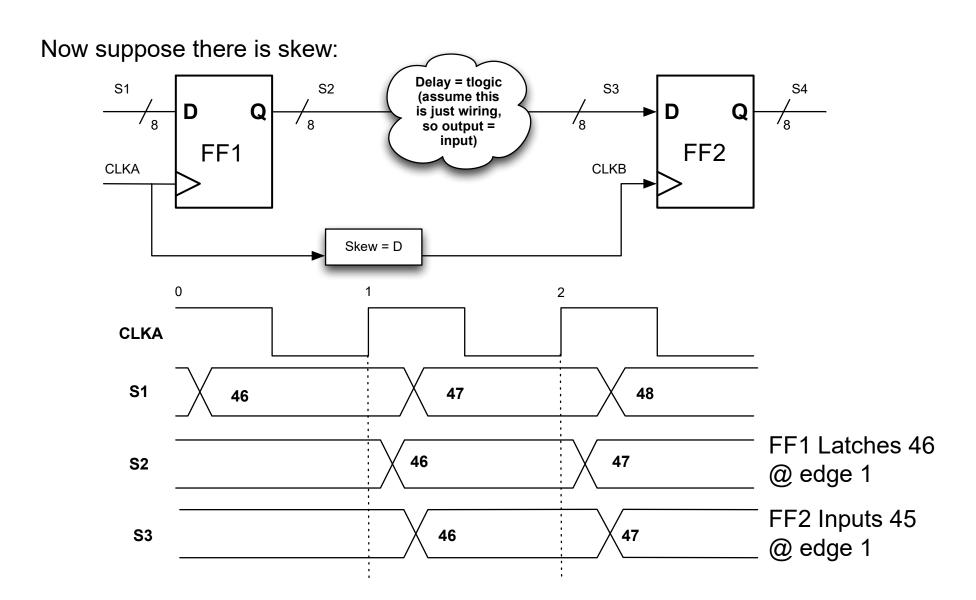
47

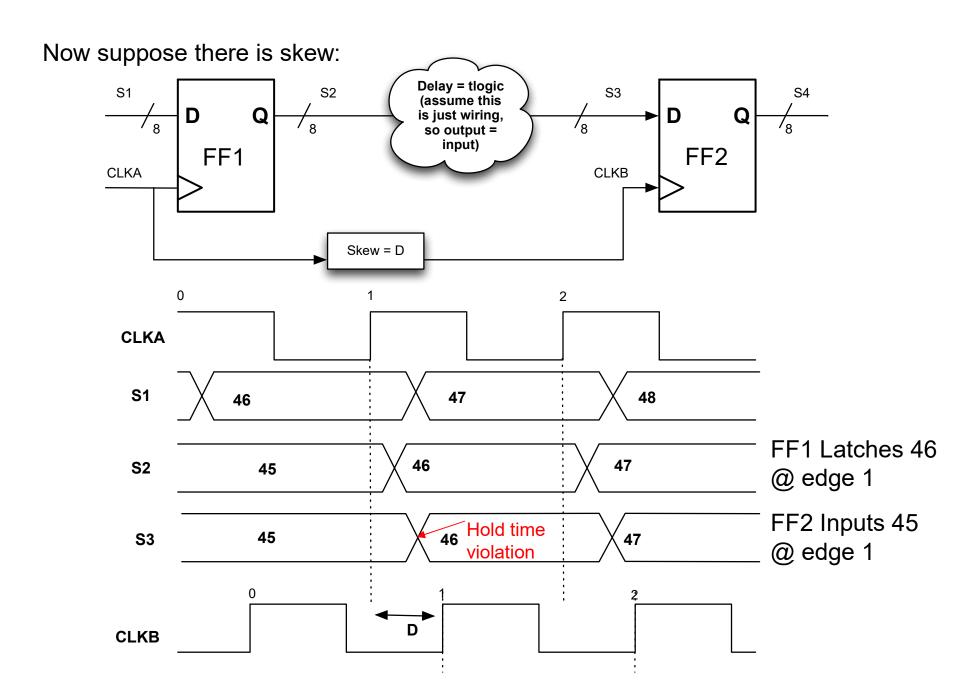


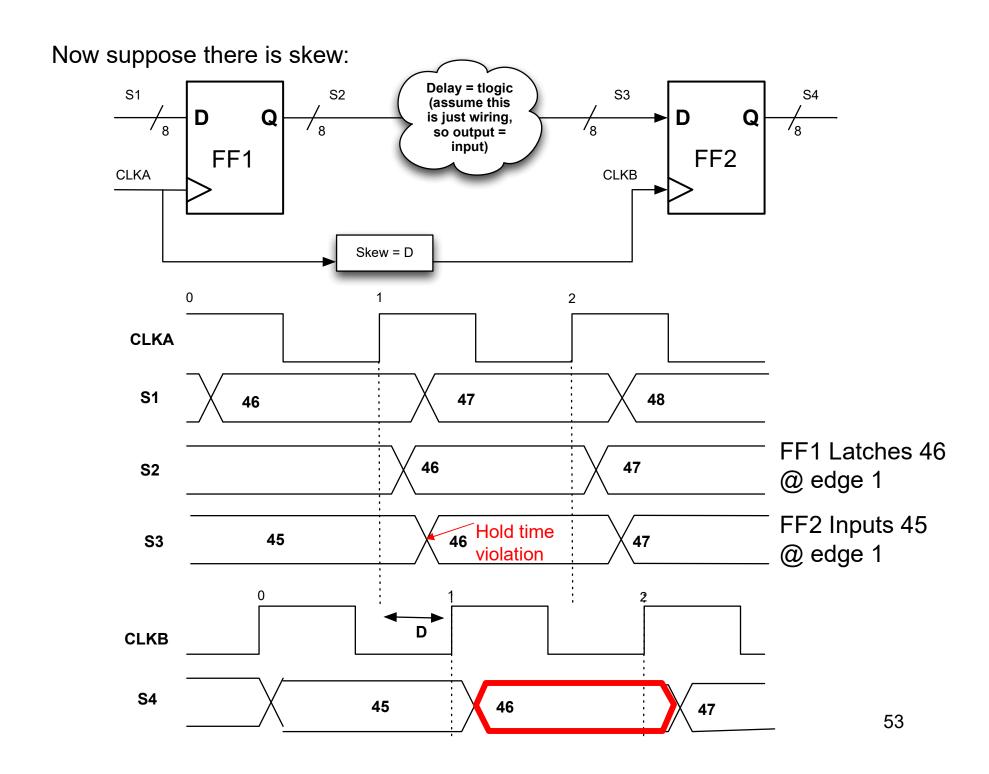
48

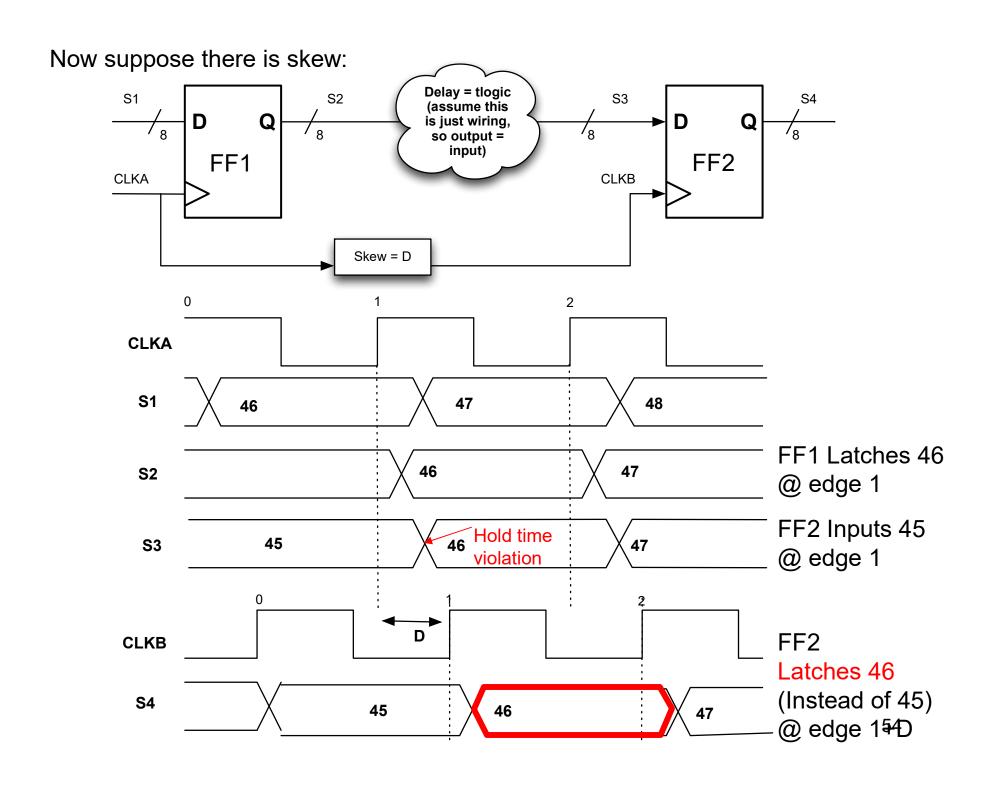










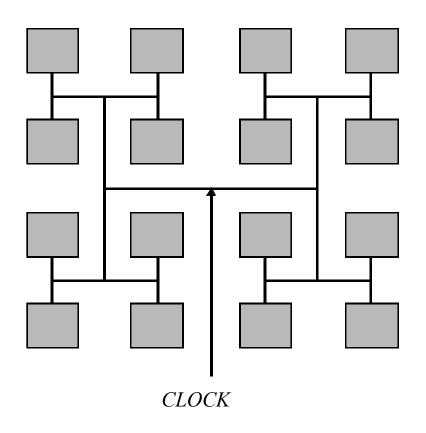


In this example, skew caused a hold violation, causing the second value "46" to race through too early. This caused a functional violation.

Slowing the clock speed doesn't help.

Lesson: Clock skew needs to be avoided as much as possible

# Clock Skew: What can we do? Custom Chip



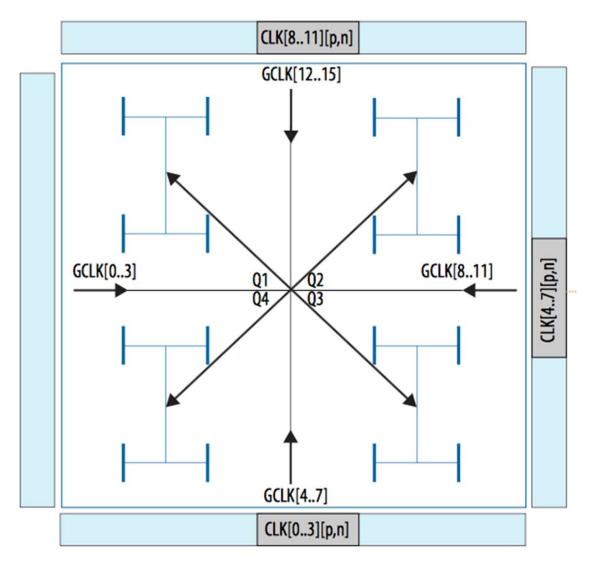
Layout clock signal to minimize skew

Need to verify timing extensively to make sure skew is not going to cause failure.

H-Tree Network

**Observe: Only Relative Skew is Important** 

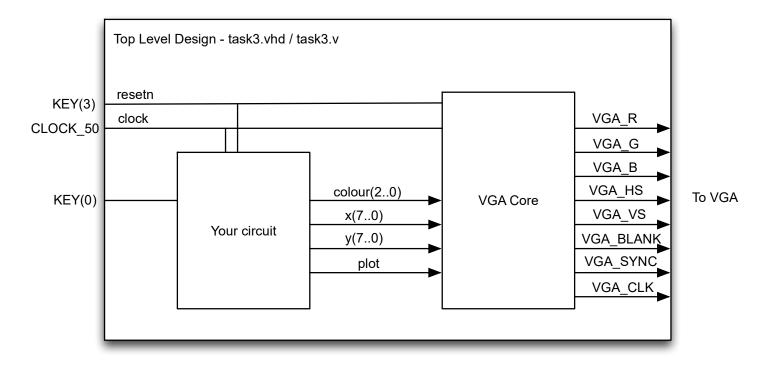
## Clock Skew: FPGAs



FPGAs have complex clock distribution networks to achieve low skew

CAD tools must verify against clock skew problems.

#### Consider the VGA core from earlier lab:



Your chip receives a 50 MHz clock

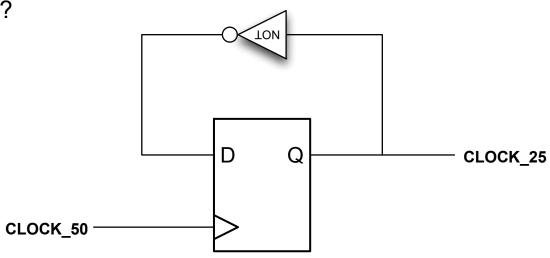
The VGA spec says VGA\_CLK needs to run at 25 MHz\*

The designer of the VGA core had to make a 25MHz clock out of a 50MHz clock

How would you make a circuit that creates a 25 MHz clock from a 50 MHz clock?

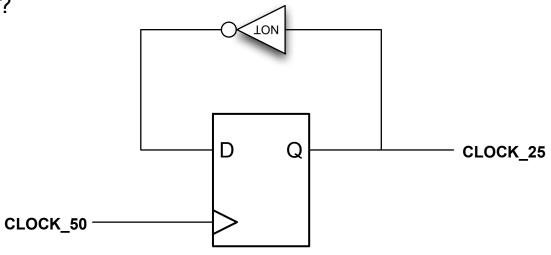
How would you make a circuit that creates a 25 MHz clock from a

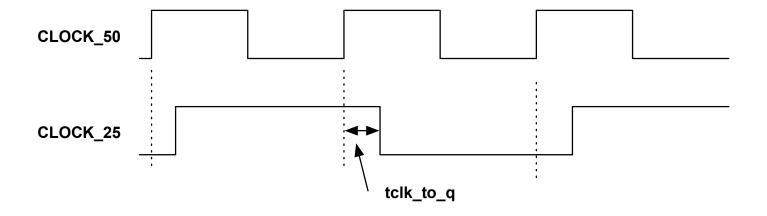
50 MHz clock?



How would you make a circuit that creates a 25 MHz clock from a

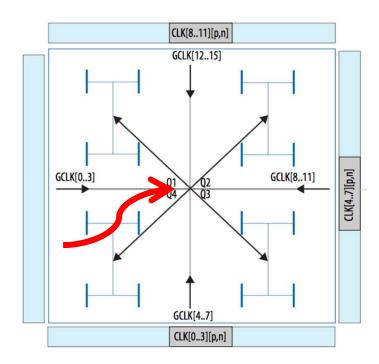
50 MHz clock?





Two problems with this solution:

- 1. Clock skew between CLOCK\_50 and CLOCK\_25
- 2. Even if you use only CLOCK\_25, it must be distributed across the entire chip. FPGA routing is slow and unpredictable.

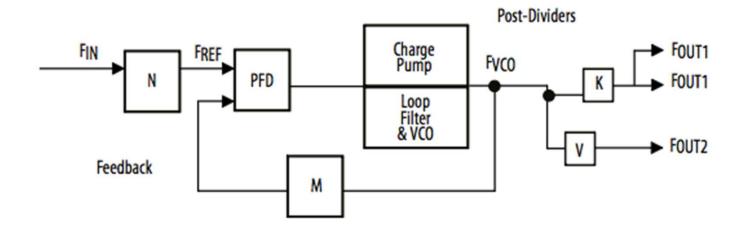


Even if CLOCK\_25 uses the global clock distribution network, we still have to route from our FF output to the source of these wires.

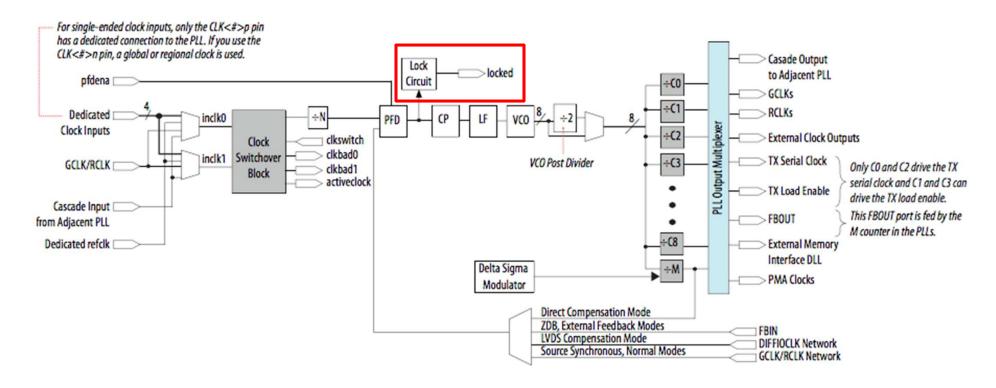
## Phase Locked Loops

Modern FPGAs contain Phase Locked Loop hard cores:

- A mixed-signal circuit ( = both analog and digital) that generates output clocks aligned to an input clock
- Can act as a clock speed divider or multiplier
- Can directly feed clock distribution network
- Automatically adjusts phase to account for clock distribution network



# Cyclone V PLL: Detail



From Cyclone V Handbook

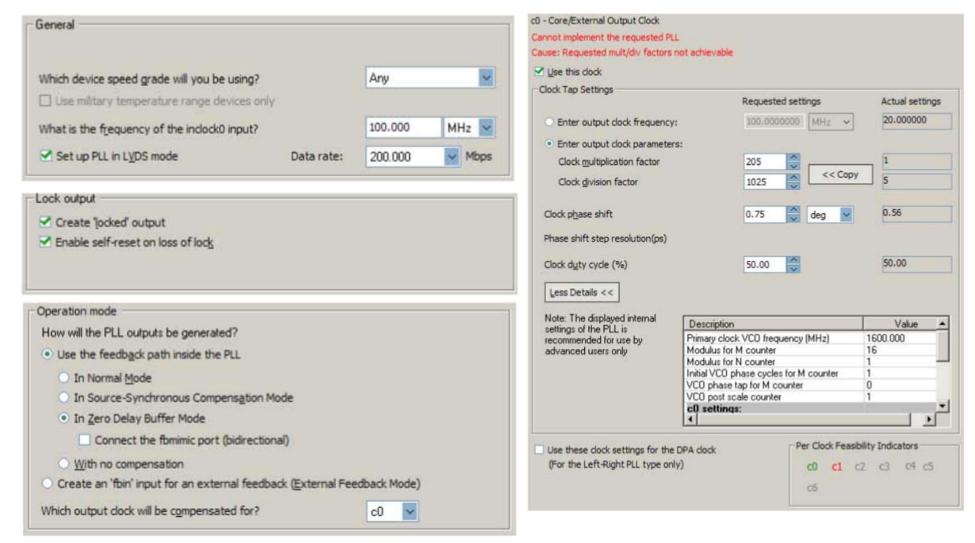
## Phase Locked Loops

If you look inside vga\_pll.v and vga\_adapter.v:

```
altpll altpll component (.inclk (clock input bus),
         .clk (clock output bus));
     defparam
        altpll component.operation mode = "NORMAL",
        altpll component.intended device family = "Cyclone
II",
        altpll component.lpm type = "altpll",
        altpll component.pll type = "FAST",
        altpll component.inclk0 input frequency = 20000,
        altpll component.primary clock = "INCLKO",
        altpll component.compensate clock = "CLKO",
        altpll component.clk0 phase shift = "0",
        altpll component.clk0 divide by = 2,
        altpll component.clk0 multiply by = 1,
vga pll mypll (CLOCK 50, clock 25);
```

Used to divide the clock frequency by 2. Can also multiply clock frequency. Don't worry about details. If you ever need it, you can look it up

# Many PLL Parameters



(parameter values shown are just examples, not recommendations)

# **Glitches**

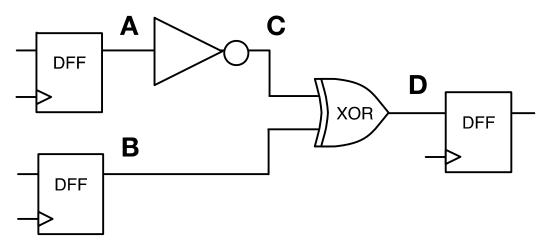
## Glitch

An undesired short-lived pulse that occurs before a signal settles to its intended value

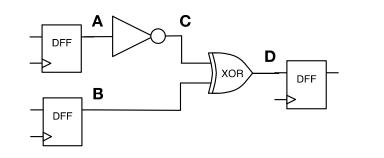
#### Causes

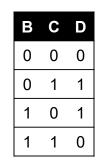
- Unequal arrival times of inputs on combinational gates
- Various electrical effects (eg crosstalk, not covered in this course)

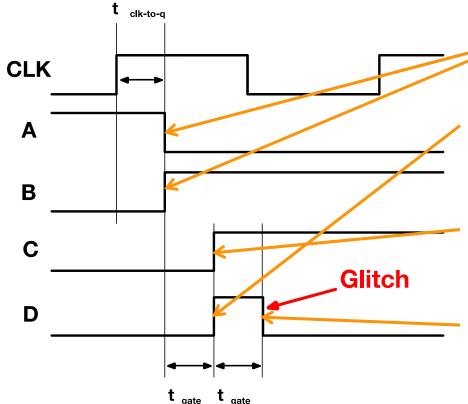
#### Example of a circuit with glitching:



## Glitch Example







- 1) A, B settle after clock-to-q delay
- 2a) B arrives immediately at XOR, causing a transition to 1 on D after one gate delay
- 2b) At the same time, A arrives at the INV immediately, causing a transition to 1 on C after one gate delay
- 3) After C settles, its value arrives at the XOR after one additional gate delay, causing a transition back to 0 on D

# Glitches – Key Take Aways

A signal may switch several times before settling to final intended value

- Source (new glitch): uneven signal arrival times on inputs
- Propagate (existing glitch): input glitch → output glitch

#### Glitches are normal ...

ok if signal is stable before anyone tries to do something with data eg, before t<sub>setup</sub> of flip-flop

#### Power and Energy

- Glitches cause extra charge/discharge cycles of output capacitance
- Unnecessary power consumption

## Learning Objectives

- 1. Understand what timing closure is and why it is difficult
- 2. Understand how pipelining can help with timing closure
- 3. Understand retiming and be able to apply it to a circuit
- 4. Be able to discuss the effects of clock skew
- 5. Understand what a PLL is used for
- 6. Understand the cause and impact of glitches caused by unequal combinational path delays